ANM.

(19) BUNDESREPUBLIK DEUTSCHLAND

DEUTSCHES PATENT- UND MARKENAMT

(12) Übersetzung der europäischen Patentschrift

(97) EP 1 016 071 B1

(10) DE 698 03 202 T 2

(51) Int. Cl.[7]: G 10 L 11/02
G 10 L 11/06

(21) Deutsches Aktenzeichen: 698 03 202.0
(86) PCT-Aktenzeichen: PCT/FR98/01979
(96) Europäisches Aktenzeichen: 98 943 998.9
(87) PCT-Veröffentlichungs-Nr.: WO 99/14737
(86) PCT-Anmeldetag: 16. 9. 1998
(87) Veröffentlichungstag der PCT-Anmeldung: 25. 3. 1999
(97) Erstveröffentlichung durch das EPA: 5. 7. 2000
(97) Veröffentlichungstag der Patenterteilung beim EPA: 16. 1. 2002
(47) Veröffentlichungstag im Patentblatt: 29. 8. 2002

(30) Unionspriorität:
9711640 18. 09. 1997 FR

(73) Patentinhaber:
Matra Nortel Communications, Quimper, FR; Eads Defence and Security Networks, Montigny Le Bretonneux, FR

(74) Vertreter:
WINTER, BRANDL, FÜRNISS, HÜBNER, RÖSS, KAISER, POLTE, Partnerschaft, 85354 Freising

(72) Erfinder:
LOCKWOOD, Philip, F-95490 Vaureal, FR; LUBIARZ, Stephane, F-95520 Osny, FR

(84) Benannte Vertragstaaten:
CH, DE, FI, GB, LI, SE

(54) VERFAHREN UND VORRICHTUNG ZUR SPRACHDETEKTION

DE 698 03 202 T 2

DE 698 03 202 T 2

BUNDESDRUCKEREI 07.02 602 350/742/3G

15.04.02

Beschreibung

Die vorliegende Erfindung betrifft digitale Verfahren zur Verarbeitung von Sprachsignalen Sie betrifft insbesondere Verfahren, die eine Erfassung von Stimmaktivität anwenden, um differenzierte Verarbeitungen je nachdem durchzuführen, ob das Signal eine Stimmaktivität aufweist oder nicht.

Die betreffenden digitalen Verfahren beziehen sich auf verschiedene Fachgebiete: Sprachcodierung für die Übertragung oder Speicherung oder Erkennung von Sprache, Verminderung von Rauschen, Echounterdrückung usw.

Die Verfahren zur Erfassung von Stimmaktivität haben als hauptsächliche Schwierigkeit die Unterscheidung zwischen der Stimmaktivität und dem sie begleitenden Lärm bzw. Rauschen. Die Zuhilfenahme eines klassischen Rauschunterdrückungsverfahrens gestattet es nicht, diese Schwierigkeit zu behandeln, da diese Verfahren wiederum Schätzungen des Rauschens anwenden, die von dem Grad der Stimmaktivität des Signals abhängen. Dieses Problem ist beispielsweise in der Schrift US-A-5659622 beschrieben.

Ein Hauptziel der vorliegenden Erfindung ist es, die Robustheit der Verfahren zur Erfassung von Stimmaktivität gegen Rauschen zu verbessern. Um dieses Ziel zu erreichen, wird ein Verfahren gemäß den Angaben in Anspruch 1 vorgeschlagen.

So schlägt die Erfindung ein Verfahren zum Erfassen von Stimmaktivität in einem in aufeinanderfolgenden Blöcken behandelten digitalen Sprachsignal vor, bei dem das Sprachsignal einer Rauschunterdrückung unter Berück-

sichtigung von Schätzungen des im Signal enthaltenen Rauschens unterzogen wird, die für jeden Block auf eine Weise aktualisiert werden, die von zumindest einem für den betreffenden Block bestimmten Grad der Stimmaktivität abhängt. Erfindungsgemäß wird eine apriorische Rauschunterdrückung des Sprachsignals eines jeden Blocks auf der Grundlage von Schätzungen des Rauschens durchgeführt, die bei der Behandlung von mindestens einem vorhergehenden Block erhalten wurden, und die Variationen der Energie des apriorisch rauschunterdrückten Signals analysiert werden, um den Grad der Stimmaktivität des Blocks zu erfassen.

Der Umstand, daß die Erfassung der Stimmaktivität (gemäß einem Verfahren, bei dem es sich im wesentlichen um jegliches bekannte Verfahren handeln kann) auf der Grundlage eines apriorisch rauschunterdrückten Signals durchgeführt wird, verbessert wesentliche die Leistungsfähigkeit dieser Erfassung, wenn das Umgebungsrauschen relativ stark ist.

In der Folge der vorliegenden Beschreibung wird das erfindungsgemäße Verfahren zum Erfassen von Stimmaktivität an einem System zur Rauschunterdrückung eines Sprachsignals veranschaulicht. Es ist ersichtlich, daß dieses Verfahren Anwendungen in zahlreichen weiteren Arten der digitalen Sprachverarbeitung finden kann, bei denen es erwünscht ist, über eine Information bezüglich des Grades der Stimmaktivität des verarbeiteten Signals zur verfügen: Codierung, Erkennung, Echounterdrückung usw.

Weitere Details und Vorteile der vorliegenden Erfindung ergeben sich aus der nachfolgenden Beschreibung von nicht-einschränkenden Ausführungsbeispielen unter Bezugnahme auf die beigefügte Zeichnung.

Es zeigt:

- Fig. 1 eine schematische Übersicht eines Rauschunterdrückungssystems, das die vorliegende Erfindung anwendet;

- Fig. 2 und 3 Organigramme von Prozeduren, die durch einen Stimmaktivitätsdetektor des Systems von Fig. 1 angewendet werden;

- Fig. 4 ein Diagramm, das die Zustände eines Automaten zur Erfassung von Stimmaktivität darstellt;

- Fig. 5 ein Diagramm zur Veranschaulichung der Variationen eines Grades der Stimmaktivität;

- Fig. 6 eine schematische Übersicht eines Moduls zur Überbewertung von Rauschen des Systems von Fig. 1;

- Fig. 7 ein Diagramm zur Veranschaulichung der Berechnung einer Maskierungskurve; und

- - Fig. 8 ein Diagramm zur Veranschaulichung der Nutzung der Maskierungskurven in dem System von Fig. 1.

Das in Fig. 1 dargestellte System zur Rauschunterdrückung behandelt ein digitales Sprachsignal s. Ein Fensterbildungsmodul 10 bringt dieses Signal s in die Form von aufeinanderfolgenden Fenstern oder Blöcken, die jeweils aus einer Anzahl N von Abtastproben eines digitalen Signals bestehen. Auf klassische Weise können diese Blöcke gegenseitige Überlappungen aufweisen. In der nachfolgenden Beschreibung wird angenommen, ohne daß dies einschränkend gedacht ist, daß die Blöcke aus N = 256 Abtastproben mit einer Abtastrate F_e von 8 kHz bestehen, mit einer Hamming-Wichtung in jedem Fenster, und Überlappungen von 50% zwischen aufeinanderfolgenden Fenstern.

Der Signalblock wird durch ein Modul 11, das einen klassischen Algorithmus der schnellen Fourier-Transformation (TFR) für die Berechnung des Moduls des Spektrums

des Signals anwendet, in den Frequenzbereich transformiert. Das Modul 11 liefert somit eine mit $S_{n,f}$ bezeichnete Gesamtheit von N = 256 Frequenzkomponenten des Sprachsignals, wobei n die Nummer des momentanen Blocks und f eine Frequenz des diskreten Spektrums bezeichnet. Aufgrund der Eigenschaften der digitalen Signale im Frequenzbereich werden nur die N/2 = 128 ersten Abtastproben verwendet.

Für die Berechnung der Schätzungen des in dem Signal s enthaltenen Rauschens wird nicht die am Ausgang der schnellen Fourier-Transformation verfügbare Frequenzauflösung verwendet, sondern eine schwächere Auflösung, die durch eine Anzahl I von Frequenzbändern bestimmt ist, welche das Band [0, $F_e/2$] des Signals abdeckt. Jedes Band i ($1 \leq i \leq I$) erstreckt sich zwischen einer unteren Frequenz f(i-1) und einer oberen Frequenz f(i), wobei f(0) = 0, und f(I) = $F_e/2$. Dieses Zerschneiden in Frequenzbänder kann gleichförmig ($f(i)-f(i-1) = F_e/2I$) sein. Es kann auch nicht gleichförmig sein (z.B. gemäß einer Barks-Skala). Ein Modul 12 berechnet die jeweiligen Mittelwerte der Spektralkomponenten $S_{n,f}$ des Sprachsignals pro Bändern, beispielsweise durch eine gleichförmige Wichtung wie etwa:

$$S_{n,i} = \frac{1}{f(i) - f(i-1)} \sum_{f \in [f(i-1), f(i)[} S_{n,f} \qquad (1)$$

Diese Mittelwertbildung vermindert die Schwankungen zwischen den Bändern durch Mitteln der Beiträge des Rauschens in diesen Bändern, wodurch die Varianz des Schätzers des Rauschens vermindert wird. Des weiteren gestattet diese Mittelwertbildung eine starke Verringerung der Komplexität des Systems.

Die gemittelten Spektralkomponenten $S_{n,i}$ werden an ein Modul 15 für die Erfassung von Stimmaktivität und an ein Modul 16 zur Schätzung des Rauschens adressiert. Diese beiden Module 15, 16 arbeiten insofern gemeinsam, als von dem Modul 15 für die verschiedenen Bänder gemessene Stimmaktivitätsgrade $\gamma_{n,i}$ von dem Modul 16 für die Schätzung der Langzeitenergie des Rauschens in den verschiedenen Bändern verwendet werden, während diese Langzeitschätzungen $\hat{B}_{n,i}$ von dem Modul 15 verwendet werden, um eine apriorische Rauschunterdrückung des Sprachsignals in den verschiedenen Bändern vorzunehmen, um die Stimmaktivitätsgrade $\gamma_{n,i}$ zu bestimmen.

Der Betrieb der Module 15 und 16 kann den in Fig. 2 und 3 dargestellten Organigrammen entsprechen.

In den Schritten 17 bis 20 führt das Modul 15 die apriorische Rauschunterdrückung des Sprachsignals in den unterschiedlichen Bändern i für den Signalblock n durch. Diese apriorische Rauschunterdrückung wird gemäß einem klassischen Vorgang zur nichtlinearen Spektralsubtraktion ausgehend von Schätzungen des Rauschens durchgeführt, welche bei einem oder mehreren vorausgegangenen Blöcken erhalten wurden. In Schritt 17 berechnet das Modul 15 mit der Auflösung der Bänder i den Frequenzgang $Hp_{n,i}$ des Filters für die apriorische Rauschunterdrückung gemäß der Formel:

$$Hp_{n,i} = \frac{S_{n,i} - \alpha'_{n-\tau1,i} \cdot \hat{B}_{n-\tau1,i}}{S_{n-\tau2,i}} \qquad (2)$$

wobei $\tau1$ und $\tau2$ als Anzahl von Blöcken ausgedrückte Verzögerungen sind ($\tau1 \geq 1$, $\tau2 \geq 0$), und $\alpha'_{n,i}$ ein Koeffizient der Überbewertung des Rauschens ist, dessen Bestimmung weiter unten erläutert wird. Die Verzögerung $\tau1$ kann

festgelegt (z.B. $\tau 1 = 1$) oder auch variabel sein. Sie ist umso geringer, je stärker man sich auf die Erfassung der Stimmaktivität verläßt.

In den Schritten 18 bis 20 werden die Spektralkomponenten $\hat{E}p_{n,i}$ berechnet gemäß:

$$\hat{E}p_{n,i} = \max\left\{Hp_{n,i} \cdot S_{n,i}, \beta p_i \cdot \hat{B}_{n-\tau 1,i}\right\} \qquad (3)$$

wobei βp_i ein Untergrenzenkoeffizient nahe 0 ist, der klassischerweise dazu dient zu vermeiden, daß das Spektrum des entrauschten Signals negative oder übermäßig schwache Werte annimmt, die ein musikalisches Geräusch hervorrufen würden.

Die Schritte 17 bis 20 bestehen somit im wesentlichen darin, von dem Spektrum des Signals eine durch den Koeffizienten $\alpha'_{n-\tau 1,i}$ majorierte Schätzung des apriorisch geschätzten Spektrums des Rauschens zu subtrahieren.

In Schritt 21 berechnet das Modul 15 die Energie des apriorisch rauschunterdrückten Signals in den verschiedenen Bändern i für den Block n: $E_{n,i} = \hat{E}p^2_{n,i}$. Es berechnet auch einen globalen Mittelwert $E_{n,0}$ der Energie des apriorisch rauschunterdrückten Signals durch eine Summe der Energien pro Band $E_{n,i}$, die mit den Breiten dieser Bänder gewichtet sind. In den nachfolgenden Angaben wird der Index $i = 0$ dazu verwendet, das globale Band des Signals zu bezeichnen.

In den Schritten 22 und 23 berechnet das Modul 15 für jedes Band i ($0 \leq i \leq I$) eine Größe $\Delta E_{n,i}$, welche für die Kurzzeitvariation der Energie des entrauschten Signals im Band i steht, sowie einen Langzeitwert $\overline{E}_{n,i}$ der Energie des entrauschten Signals im Band i. Die Größe $\Delta E_{n,i}$ kann

berechnet werden durch eine vereinfachte Ableitungsformel: $\Delta E_{n,i} = \left| \frac{E_{n-4,i} + E_{n-3,i} - E_{n-1,i} - E_{n,i}}{10} \right|$. Was die Langzeitenergie $\overline{E}_{n,i}$ betrifft, so kann diese mit Hilfe eines Vergessensfaktors B1 wie etwa $0 < B1 < 1$ berechnet werden, nämlich $\overline{E}_{n,i} = B1 \cdot \overline{E}_{n-1,i} + (1 - B1) \cdot E_{n,i}$.

Nach der Berechnung der Energien $E_{n,i}$ des rauschunterdrückten Signals, seiner Kurzzeitvariationen $\Delta E_{n,i}$ und seiner Langzeitwerte $\overline{E}_{n,i}$ auf die in Fig. 2 angegebene Weise berechnet das Modul 15 für jedes Band i ($0 \leq i \leq I$) einen Wert ρ_i, der für die Evolution der Energie des rauschunterdrückten Signals steht. Diese Berechnung wird in den Schritten 25 bis 36 von Fig. 3 vorgenommen, die für jedes Band i zwischen i=0 und i=I durchgeführt werden. Diese Berechnung wendet einen Langzeitschätzer ba_i der Umhüllenden des Rauschens, einen internen Schätzer bi_i und einen Zähler b_i für verrauschte Blöcke an.

In Schritt 25 wird die Größe $\Delta E_{n,i}$ mit einem Schwellwert $\varepsilon 1$ verglichen. Wenn der Schwellwert $\varepsilon 1$ nicht erreicht wird, wird der Zähler b_i in Schritt 26 um eine Einheit inkrementiert. In Schritt 27 wird der Langzeitschätzer ba_i mit dem Wert der geglätteten Energie $\overline{E}_{n,i}$ verglichen. Falls $ba_i \geq \overline{E}_{n,i}$, wird der Schätzer ba_i gleich dem geglätteten Wert $\overline{E}_{n,i}$ in Schritt 26 genommen, und der Zähler b_i wird auf Null zurückgesetzt. Die Größe ρ_i, die gleich dem Verhältnis $ba_i/\overline{E}_{n,i}$ genommen wird (Schritt 36), ist somit gleich 1.

Wenn Schritt 27 ergibt, daß $ba_i < \overline{E}_{n,i}$, wird der Zähler b_i in Schritt 29 mit einem Grenzwert bmax verglichen. Falls $b_i >$ bmax, wird angenommen, daß das Signal zu stationär ist, um Stimmaktivität zu unterstützen. Daraufhin wird der oben genannte Schritt 28 durchgeführt, der in der Annahme besteht, daß der Block nur Rauschen

beinhaltet. Falls $b_i \leq bmax$ in Schritt 29, wird der interne Schätzer bi_i in Schritt 33 berechnet gemäß:

$$bi_i = (1-Bm) \cdot \bar{E}_{n,i} + Bm \cdot ba_i \qquad (4)$$

In dieser Formel steht Bm für einen zwischen 0,90 und 1 liegenden Aktualisierungskoeffizienten. Sein Wert ist je nach dem Zustand eines Automaten für die Erfassung von Stimmaktivität verschieden (Schritte 30 bis 32). Dieser Zusand δ_{n-1} ist derjenige, der bei der Verarbeitung des vorherigen Blockes bestimmt wurde. Falls sich der Automat in einem Zustand der Erfassung von Sprache befindet ($\delta_{n-1} = 2$ in Schritt 30), nimmt der Koeffizient Bm einen Wert Bmp an, der sehr nahe bei 1 liegt, damit der Schätzer des Rauschens bei Vorhandensein von Sprache sehr geringfügig aktualisiert wird. Im entgegengesetzten Fall nimmt der Koeffizient Bm einen geringeren Wert Bms an, um in einer Stillephase eine bedeutendere Aktualisierung des Schätzers des Rauschens zu ermöglichen. In Schritt 34 wird der Abstand $ba_i - bi_i$ zwischen dem Langzeitschätzer und dem internen Schätzer des Rauschens mit einem Schwellwert $\varepsilon 2$ verglichen. Wenn der Schwellwert $\varepsilon 2$ nicht erreicht wird, wird der Langzeitschätzer ba_i in Schritt 35 mit dem Wert des internen Schätzers bi_i aktualisiert. Andernfalls bleibt der Langzeitschätzer ba_i unverändert. Es wird somit vermieden, daß abrupte Variationen aufgrund eines Sprachsignals zu einer Aktualisierung des Schätzers des Rauschens führen.

Nach dem Erhalt der Größen ρ_i nimmt das Modul 15 die Entscheidungen der Stimmaktivität in Schritt 37 vor. Das Modul 15 aktualisiert zuerst den Zustand des Erfassungsautomaten gemäß der für die Gesamtheit des Bandes des Signals berechneten Größe ρ_0. Der neue Zustand δ_n des

Automaten hängt von dem vorhergegangenen Zustand δ_{n-1} und von ρ_0 ab, wie in Fig. 4 dargestellt ist.

Vier Zustände sind möglich: $\delta = 0$ erfaßt Stille bzw. Abwesenheit von Sprache; $\delta = 2$ erfaßt das Vorhandensein einer Stimmaktivität; und die Zustände $\delta = 1$ und $\delta = 3$ sind dazwischenliegende Zustände des Anstiegs und Abfallens. Wenn sich der Automat im Zustand von Stille ($\delta_{n-1} = 0$) befindet, bleibt er dort, wenn ρ_0 nicht eine erste Schwelle SE1 übersteigt, und geht im entgegengesetzten Fall in den Anstiegszustand über. Im Anstiegszustand ($\delta_{n-1} = 1$) kehrt er in den Zustand von Stille zurück, wenn ρ_0 kleiner als der Schwellwert SE1 ist, geht in den Zustand der Sprache über, wenn ρ_0 größer als eine über der Schwelle SE1 liegende Schwelle SE2 ist, und bleibt im Anstiegszustand, falls $SE1 \leq \rho_0 \leq SE2$. Wenn sich der Automat im Zustand der Sprache ($\delta_{n-1} = 2$) befindet, so bleibt er dort, falls ρ_0 eine unter der Schwelle SE2 liegende dritte Schwelle SE3 ist, und geht im entgegengesetzten Fall in den Abstiegszustand über. Im Abstiegszustand ($\delta_{n-1} = 3$) kehrt der Automat in den Zustand der Sprache zurück, falls de ρ_0 größer als der Schwellwert SE2 ist, kehrt in den Zustand der Stille zurück, wenn ρ_0 diesseits eines unter dem Schwellwert SE2 liegenden vierten Schwellwerts SE4 ist, und bleibt im Abstiegszustand, falls $SE4 \leq \rho_0 \leq SE2$.

In Schritt 37 berechnet das Modul 15 des weiteren die Stimmaktivitätsgrade $\gamma_{n,i}$ in jedem Band $i \geq 1$. Dieser Grad $\gamma_{n,i}$ ist vorzugsweise ein nicht-binärer Parameter, d.h. die Funktion $\gamma_{n,i} = g(\rho_i)$ ist eine Funktion, die in Abhängigkeit von den durch die Größe ρ_i angenommenen Werten kontinuierlich zwischen 0 und 1 variiert. Diese Funktion besitzt beispielsweise den in Fig. 5 dargestellten Verlauf.

Das Modul 16 berechnet die Schätzungen des Rauschens pro Band, die im Rauschunterdrückungsvorgang verwendet werden, unter Anwendung der aufeinanderfolgenden Werte der Komponenten $S_{n,i}$ und der Stimmaktivitätsgrade $\gamma_{n,i}$. Dies entspricht den Schritten 40 bis 42 von Fig. 3. In Schritt 40 wird bestimmt, ob der Automat für die Erfassung von Stimmaktivität aus dem Anstiegszustand in den Zustand der Sprache übergegangen ist. Falls ja, werden die vorausgehend für jedes Band $i \geq 1$ berechneten beiden letzten Schätzungen $\hat{B}_{n-1,i}$ und $\hat{B}_{n-2,i}$ gemäß dem vorausgegangenen Schätzwert $\hat{B}_{n-3,i}$ korrigiert. Diese Korrektur wird durchgeführt, um den Umstand zu berücksichtigen, daß in der Anstiegsphase ($\delta = 1$) die Langzeitschätzungen der Energie des Rauschens in dem Vorgang für die Erfassung von Stimmaktivität (Schritte 30 bis 33) so berechnet werden konnten, als ob das Signal nur Rauschen beinhaltete (Bm = Bms), so daß die Gefahr besteht, daß sie mit einem Fehler behaftet sind.

In Schritt 42 aktualisiert das Modul 16 die Schätzungen des Rauschens pro Band gemäß den Formeln:

$$\bar{B}_{n,i} = \lambda_B \cdot \hat{B}_{n-1,i} + (1 - \lambda_B) \cdot S_{n,i} \qquad (5)$$

$$\hat{B}_{n,i} = \gamma_{n,i} \cdot \hat{B}_{n-1,i} + (1 - \gamma_{n,i}) \cdot \bar{B}_{n,i} \qquad (6)$$

wobei λ_B einen Vergessensfaktor wie etwa $0 < \lambda_B < 1$ bezeichnet. Formel (6) zeigt die Berücksichtigung des nicht-binären Stimmaktivitätsgrades $\gamma_{n,i}$.

Wie obenstehend angegeben wurde, sind die Langzeitschätzungen des Rauschens $\hat{B}_{n,i}$ Gegenstand einer Überbewertung durch ein Modul 45 (Fig. 1), bevor die Rauschunterdrückung mittels nichtlinearer Spektralsubtraktion

vorgenommen wird. Das Modul 45 berechnet den oben genannten Koeffizienten der Überbewertung $\alpha'_{n,i}$ sowie eine majorierte Schätzung $\hat{B}'_{n,i}$ die im wesentlichen $\alpha'_{n,i} \cdot \hat{B}'_{n,i}$ entspricht.

Die Strukturierung des Überbewertungsmoduls 45 ist in Fig. 6 dargestellt. Die majorierte Schätzung $\hat{B}'_{n,i}$ wird erhalten durch Kombinieren der Langzeitschätzung $\hat{B}_{n,i}$ und eines Maßes $\Delta B^{max}_{n,i}$ der Veränderlichkeit der Rauschkomponente in dem Band i um seine Langzeitschätzung. Bei dem betrachteten Beispiel ist dieses Kombinieren im wesentlichen eine einfache Summe, die von einem Addierer 46 erstellt wird. Es könnte sich hierbei auch um eine gewichtete Summe handeln.

Der Überbewertungskoeffizient $\alpha'_{n,i}$ ist gleich dem Verhältnis zwischen der vom Addierer 46 gelieferten Summe $\hat{B}_{n,i} + \Delta B^{max}_{n,i}$ und der verzögerten Langzeitschätzung $\hat{B}_{n-\tau 3,i}$ (Teiler 47), die nach oben hin durch einen Grenzwert α_{max} beschränkt ist, beispielsweise $\alpha_{max} = 4$ (Block 48). Die Verzögerung $\tau 3$ dient gegebenenfalls dazu, in den Anstiegsphasen ($\delta = 1$) den Wert des Überbewertungskoeffizienten $\alpha'_{n,i}$ zu korrigieren, bevor die Langzeitschätzungen durch die Schritte 40 und 41 von Fig. 3 korrigiert worden sind (z.B. $\tau 3 = 3$).

Die majorierte Schätzung $\hat{B}'_{n,i}$ wird schließlich gleich $\alpha'_{n,i} \cdot \hat{B}'_{n-\tau 3,i}$ genommen (Multiplizierer 49).

Das Maß $\Delta B^{max}_{n,i}$ der Veränderlichkeit des Rauschens reflektiert die Varianz des Schätzers des Rauschens. Es wird in Abhängigkeit von den Werten von $S_{n,i}$ und von $\hat{B}_{n,i}$ für eine bestimmte Anzahl von vorherigen Blöcken berechnet, an denen das Sprachsignal keine Stimmaktivität in dem Band i aufweist. Es ist eine Funktion der für eine

Anzahl K von Blöcken mit Stille ($n-k \leq n$) berechneten Abstände $|S_{n-k,i} - \hat{B}_{n-k,i}|$. In dem dargestellten Beispiel ist diese Funktion einfach das Maximum (Block 50). Für jeden Block n wird der Grad der Stimmaktivität $\gamma_{n,i}$ mit einem Schwellwert (Block 51) verglichen, um zu entscheiden, ob der in 52-53 berechnete Abstand $|S_{n,i} - \hat{B}_{n,i}|$ in eine Warteschlange 54 mit K Stellen geladen werden muß, die im Ersteingang/Erstausgang-Modus (FIFO) organisiert ist. Falls $\gamma_{n,i}$ den Schwellwert nicht übersteigt (der gleich 0 sein kann, falls die Fuunktion g() die Form von Fig. 5 besitzt), wird die FIFO nicht versorgt, während sie es im entgegengesetzten Fall wird. Der in der FIFO 54 enthaltene Maximalwert wird dann als Maß $\Delta B_{n,i}^{max}$ der Veränderlichkeit geliefert.

Das Maß $\Delta B_{n,i}^{max}$ der Veränderlichkeit kann als Variante in Abhängigkeit von den Werten $S_{n,f}$ (anstatt $S_{n,i}$) und $\hat{B}_{n,i}$ erhalten werden. Anschließend wird auf die gleiche Weise, mit der Ausnahme, daß die FIFO 54 $|S_{n-k,i} - \hat{B}_{n-k,i}|$ nicht enthält, vorgegangen, jedoch eher $\max_{f \in [f(i-1),f(i)[} |S_{n-k,f} - \hat{B}_{n-k,i}|$.

Aufgrund der unabhängigen Langzeitschätzungen der Schwankungen des Rauschens $\hat{B}_{n,i}$ und seiner Kurzzeitveränderlichkeit $\Delta B_{n,i}^{max}$ stellt der majorierte Schätzer $\hat{B}'_{n,i}$ eine ausgezeichnete Robustheit des Rauschunterdrückungsverfahrens gegen musikalische Geräusche zur Verfügung.

Eine erste Phase der spektralen Subtraktion wird durch das in Fig. 1 dargestellte Modul 55 verwirklicht. Diese Phase liefert vor der Auflösung der Bänder i ($1 \leq i \leq I$) den Frequenzgang $H_{n,i}^{I}$ eines ersten Rauschunterdrückungsfilters in Abhängigkeit von den Komponenten $S_{n,i}$ und $\hat{B}_{n,i}$ und den Überbewertungskoeffizienten $\alpha'_{n,i}$. Diese

Berechnung kann für jedes Band i durchgeführt werden gemäß der Formel:

$$H^1_{n,i} = \frac{\max\left\{S_{n,i} - \alpha'_{n,i} \cdot \hat{B}_{n,i}, \beta^1_i \cdot \hat{B}_{n,i}\right\}}{S_{n-\tau4,i}} \qquad (7)$$

wobei $\tau4$ eine als $\tau4 \geq 0$ (z.B. $\tau4 = 0$) bestimmte ganzzahlige Verzögerung ist. In dem Ausdruck (7) stellt der Koeffizient β^1_i wie der Koeffizient βp_i der Formel (3) eine Untergrenze dar, die klassischerweise zur Vermeidung von negativen oder zu kleinen Werten des rauschunterdrückten Signals dient.

Auf bekannte Weise (EP 0 534 837) könnte der Überbewertungskoeffizient $\alpha'_{n,i}$ in der Formel (7) durch einen anderen Koeffizienten ersetzt werden, der gleich einer Funktion von $\alpha'_{n,i}$ und einer Schätzung des Rauschabstandes (z.B. $S_{n,i}/\hat{B}_{n,i}$) ist, wobei diese Funktion gemäß dem Schätzwert des Rauschabstandes abnehmend ist. Diese Funktion ist somit gleich $\alpha'_{n,i}$ für die kleinsten Werte des Rauschabstandes. Wenn das Signal stark verrauscht ist, ist es nämlich a priori nicht sinnvoll, den Überbewertungsfaktor zu vermindern. Vorteilhaft nimmt diese Funktion für die höchsten Werte des Rauschabstandes gegen Null hin ab. Dies ermöglicht einen Schutz der energiereichsten Zonen des Spektrums, in denen das Sprachsignal am bedeutendsten ist, wobei die von dem Signal zu subtrahierende Größe somit gegen Null tendiert.

Diese Strategie kann verfeinert werden, indem sie selektiv auf die Harmonischen der Tonfrequenz ("pitch") des Sprachsignals angewendet wird, wenn dieses eine Stimmaktivität aufweist.

Somit wird bei der in Fig. 1 dargestellten Ausführungsform eine zweite Phase der Rauschunterdrückung durch ein Modul 56 zum Schutz der Harmonischen durchgeführt. Dieses Modul berechnet mit der Auflösung der Fourier-Transformierung den Frequenzgang $H^2_{n,f}$ eines zweiten Rauschunterdrückungsfilters in Abhängigkeit von den Parametern $H^1_{n,i}$, $\alpha'_{n,i}$, $\hat{B}_{n,i}$, δ_n, $S_{n,i}$ und der außerhalb der Stillephasen durch ein Modul für die harmonische Analyse 57 berechneten Tonfrequenz $f_p = F_e/T_p$. In einer Stillephase ($\delta_n = 0$) ist das Modul 56 nicht in Betrieb, d.h. $H^2_{n,f} = H^1_{n,i}$ für jede Frequenz f eines Bandes i. Das Modul 57 kann jegliches bekannte Verfahren für die Analyse des Sprachsignals des Blocks anwenden, um die Periode T_p zu bestimmen, die als ganze Zahl oder Bruchteil von Abtastproben angegeben wird, z.B. ein lineares Prädiktionsverfahren.

Der durch das Modul 56 zur Verfügung gestellte Schutz kann darin bestehen, daß für jede zu einem Band i gehörige Frequenz f durchgeführt wird:

$$\begin{cases} H^2_{n,f} = 1 \text{ falls} \begin{cases} S_{n,i} - \alpha'_{n,i} \cdot \hat{B}_{n,i} > \beta^2_i \cdot \hat{B}_{n,i} & (8) \\ \text{und ganzzahliges } \exists\eta \;/\; |f - \eta \cdot f_p| \le \Delta f \,/\, 2 & (9) \end{cases} \\ \text{andernfalls } H^2_{n,f} = H^1_{n,f} \end{cases}$$

$\Delta f = F_e/N$ stellt die spektrale Auflösung der Fourier-Transformation dar. Wenn $H^2_{n,f} = 1$, ist die von der Komponente $S_{n,f}$ zu substrahierende Größe Null. In dieser Berechnung drücken die Untergrenzenkoeffizienten β^2_i (z.B. $\beta^2_i = \beta^1_i$) den Umstand aus, daß bestimmte Harmonische der Tonfrequenz f_p von Rauschen maskiert sein können, so daß es nicht sinnvoll ist, sie zu schützen.

Diese Schutzstrategie wird vorzugsweise für jede der Frequenzen angewendet, die am nächsten zu den Harmonischen von f_p sind, d.h. auf jedes ganzzahlige η.

Wenn man mit δf_p die Frequenzauflösung bezeichnet, bei der das Analysemodul 57 die geschätzte Tonfrequenz f_p erzeugt, d.h. daß die reelle Tonfrequenz zwischen $f_p - \delta f_p/2$ und $f_p + \delta f_p/2$ liegt, dann kann der Abstand zwischen der η-ten Harmonischen der reellen Tonfrequenz und ihrer Schätzung $\eta \times f_p$ (Bedingung (9)) bis $\pm\eta \times \delta f_p/2$ gehen. Bei hohen Werten von η kann dieser Abstand größer als die halbe spektrale Auflösung $\Delta f/2$ der Fourier-Transformierten sein. Um diese Unsicherheit zu berücksichtigen und einen guten Schutz der Harmonischen der reellen Tonfrequenz zu gewährleisten, kann jede der Frequenzen des Intervalls $\left[\eta \times f_p - \eta \times \delta f_p / 2, \eta \times f_p + \eta \times \delta f_p / 2\right]$ geschützt werden, d.h. die obenstehende Bedingung (9) kann ersetzt werden durch:

$$\text{ganzzahliges } \exists\eta \,/\, \left|f - \eta \cdot f_p\right| \leq \left(\eta \cdot \delta f_p + \Delta f\right) / 2 \quad (9')$$

Diese Schutzart (Bedingung 9') ist von besonderem Interesse, wenn die Werte von η groß sein können, insbesondere falls das Verfahren in einem Breitbandsystem verwendet wird.

Für jede geschützte Frequenz kann der korrigierte Frequenzgang $H^2_{n,f}$ gemäß der obenstehenden Angabe gleich 1 sein, was der Substraktion einer Größe Null im Rahmen der spektralen Subtraktion entspricht, d.h. einem kompletten Schutz der betreffenden Frequenz. Allgemeiner gesagt, dieser korrigierte Frequenzgang $H^2_{n,f}$ könnte je nach dem gewünschten Schutzgrad gleich einem zwischen 1 und $H^1_{n,f}$ liegenden Wert genommen werden, was der Subtraktion einer Größe entspricht, die kleiner als diejenige ist, die zu

subtrahieren wäre, wenn die betreffende Frequenz nicht geschützt wäre.

Die Spektralkomponenten $S^2_{n,f}$ eines rauschunterdrückten Signals werden durch einen Multiplizierer 58 berechnet:

$$S^2_{n,f} = H^2_{n,f} \cdot S_{n,f} \qquad (10)$$

Dieses Signal $S^2_{n,f}$ wird an ein Modul 60 geliefert, das für jeden Block n eine Maskierungskurve berechnet durch Anwenden eines psychoakustischen Modells der Gehörwahrnehmung durch das menschliche Ohr.

Das Phänomen der Maskierung ist ein von der Funktion des menschlichen Ohrs her bekanntes Prinzip. Wenn zwei Frequenzen gleichzeitig gehört werden, ist es möglich, daß eine von den beiden nicht mehr hörbar ist. Man sagt dann, daß diese maskiert ist.

Es gibt verschiedene Verfahrensweisen für die Berechnung der Maskierungskurven. Beispielsweise kann die von J.D. Johnston ("Transform Coding of Audio Signals Using Perceptual Noise Criteria", IEEE Journal on Selected Area in Communications, Vol. 6, Nr. 2, Februar 1988) entwickelte angewendet werden. Bei dieser Verfahrensweise wird in der Frequenzskala der Barks gearbeitet. Die Maskierungskurve wird als die Faltung der Funktion der spektralen Dehnung der Basilarmembran im Bark-Bereich mit dem anregenden Signal betrachtet, bestehend in der vorliegenden Anwendung aus dem Signal $S^2_{n,f}$. Die spektrale Dehnungsfunktion kann auf die in Fig. 7 dargestellte Weise modelliert werden. Für jedes Bark-Band wird der Beitrag der in Betracht gezogenen niederen und hohen Bänder durch die Funktion der Dehnung der Basilarmembran berechnet:

$$C_{n,q} = \sum_{q'=0}^{q-1} \frac{S^2_{n,q'}}{\left(10^{10/10}\right)^{(q-q')}} + \sum_{q'=q+1}^{Q} \frac{S^2_{n,q'}}{\left(10^{25/10}\right)^{(q'-q)}} \quad (11)$$

wobei die Indices q und q' die Bark-Bänder ($0 \leq q,q' \leq Q$) bezeichnen, und $S^2_{n,q'}$ für den Mittelwert der Komponenten $S^2_{n,f}$ des rauschunterdrückten Anregungssignals für die diskreten Frequenzen f steht, die zum Bark-Band q' gehören.

Der Maskierungsschwellwert $M_{n,q}$ wird erhalten durch das Modul 60 für Bark-Band q gemäß der Formel:

$$M_{n,q} = C_{n,q}/R_q \quad (12)$$

in der R_q von dem mehr oder minder stimmhaften Charakter des Signals abhängt. Auf bekannte Weise ist eine mögliche Form von R_q:

$$10.\log_{10}(R_q) = (A+q) \cdot \chi + B \cdot (1 - \chi) \quad (13)$$

wobei A = 14,5 und B = 5,5. χ bezeichnet einen Stimmhaftigkeitsgrad des Sprachsignals, der zwischen Null (keine Stimmhaftigkeit) und 1 (stark stimmhaftes Signal) variiert. Der Parameter χ kann die bekannte Form aufweisen:

$$\chi = \min\left\{\frac{SFM}{SFM_{max}}, 1\right\} \quad (12)$$

wobei SFM in Dezibel das Verhältnis zwischen dem arithmetischen Mittel und dem geometrischen Mittel der Energie der Bark-Bänder angibt, und $SFM_{max} = -60$ dB.

Das Rauschunterdrückungssystem weist darüber hinaus ein Modul 62 auf, das den Frequenzgang des Rauschunterdrückungsfilters in Abhängigkeit von der durch das Modul 60 berechneten Maskierungskurve $M_{n,q}$ und den durch das Modul 45 berechneten majorierten Schätzungen korrigiert. Das Modul 62 entscheidet über das Rauschunterdrückungsniveau, das tatsächlich erzielt werden soll.

Durch einen Vergleich der Umhüllenden der majorierten Schätzung des Rauschens mit der durch die Maskierungsschwellwerte $M_{n,q}$ gebildeten Umhüllenden wird entschieden, das Signal nur in dem Maße zu entrauschen, in dem die majorierte Schätzung $\hat{B}'_{n,i}$ die Maskierungskurve übersteigt. Dies vermeidet eine nutzlose Unterdrückung von durch Sprache maskiertem Rauschen.

Die neue Antwort $H^3_{n,f}$ für eine zu dem Band i gehörende Frequenz f, die durch das Modul 12 und im Bark-Band q definiert wird, hängt somit von dem relativen Abstand zwischen der majorierten Schätzung $\hat{B}'_{n,i}$ der entsprechenden Spektralkomponente des Rauschens und der Maskierungskurve $M_{n,q}$, folgendermaßen ab:

$$H^3_{n,f} = 1 - \left(1 - H^2_{n,f}\right) \cdot \max\left\{\frac{\hat{B}'_{n,i} - M_{n,q}}{\hat{B}'_{n,i}}, 0\right\} \qquad (14)$$

Anders ausgedrückt, die bei dem Vorgang der Spektralsubtraktion mit dem Frequenzgang $H^3_{n,f}$ von einer Spektralkomponente $S_{n,f}$ zu subtrahierende Größe ist im wesentlichen gleich dem Minimum zwischen der bei dem Vorgang der Spektralsubtraktion mit dem Frequenzgang $H^2_{n,f}$ von dieser Spektralkomponente zu subtrahierenden Größe einerseits und dem Bruchteil der majorierten Schätzung $\hat{B}'_{n,i}$ der entsprechenden Spektralkomponente des Rauschens anderer-

seits, die gegebenenfalls die Maskierungskurve $M_{n,q}$ übersteigt.

Fig. 8 veranschaulicht das Prinzip der durch das Modul 62 angewendeten Korrektur. Sie zeigt schematisch ein Beispiel für eine auf der Grundlage der Spektralkomponenten $S^2_{n,f}$ des rauschunterdrückten Signals sowie der majorierten Schätzung $\hat{B}'_{n,i}$ des Spektrums des Rauschens berechnete Maskierungskurve $M_{n,q}$. Die schließlich von den Komponenten $S_{n,f}$ zu subtrahierende Größe ist die durch die schraffierten Bereiche dargestellte, d.h. diejenige, die auf den Bruchteil der majorierten Schätzung $\hat{B}'_{n,i}$ der Spektralkomponenten des Rauschens, das die Maskierungskurve übersteigt, begrenzt ist.

Diese Subtraktion wird durch Multiplizieren des Frequenzgangs $H^3_{n,f}$ des Rauschunterdrückungsfilters mit den Spektralkomponenten $S_{n,f}$ des Sprachsignals (Multiplizierer 64) durchgeführt. Ein Modul 65 rekonstruiert sodann das rauschunterdrückte Signal im Zeitbereich mittels Durchführung der schnellen inversen Fourier-Transformierung (TFRI) der vom Multiplizierer 64 gelieferten Abtastproben der Frequenz $S^3_{n,f}$. Bei jedem Block werden einzig die N/2 = 128 ersten Abtastproben des durch das Modul 65 erzeugten Signals als endgültiges rauschunterdrücktes Signal s^3 geliefert, nach Rekonstruktion mittels Addition-Überlappung mit den N/2 = 128 letzten Abtastproben des vorangegangenen Blocks (Modul 66).

15.04.02

Ansprüche

1. Verfahren zum Erfassen von Stimmaktivität in einem in aufeinanderfolgenden Blöcken behandelten digitalen Sprachsignal (s), bei dem das Sprachsignal einer Rauschunterdrückung unter Berücksichtigung von Schätzungen des im Signal enthaltenen Rauschens unterzogen wird, die für jeden Block auf eine Weise aktualisiert werden, die von zumindest einem für den betreffenden Block bestimmten Grad der Stimmaktivität ($\gamma_{n,i}$) abhängt, dadurch gekennzeichnet, daß eine apriorische Rauschunterdrückung des Sprachsignals eines jeden Blocks auf der Grundlage von Schätzungen des Rauschens durchgeführt wird, die bei der Behandlung von mindestens einem vorhergehenden Block erhalten wurden, und die Variationen der Energie des apriorisch rauschunterdrückten Signals analysiert werden, um den Grad der Stimmaktivität des Blocks zu erfassen.

2. Verfahren nach Anspruch 1, bei dem der Grad der Stimmaktivität ($\gamma_{n,i}$) ein nicht-binärer Parameter ist.

3. Verfahren nach Anspruch 2, bei dem der Grad der Stimmaktivität ($\gamma_{n,i}$) eine ständig zwischen 0 und 1 variierende Funktion ist.

4. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die Schätzungen des Rauschens in verschiedenen Frequenzbändern des Signals erhalten werden, die apriorische Rauschunterdrückung Band für Band durchgeführt wird, und ein Grad der Stimmaktivität ($\gamma_{n,i}$) für jedes Band bestimmt wird.

5. Verfahren nach einem der vorhergehenden Ansprüche, bei dem eine Schätzung des Rauschens $\hat{B}_{n,i}$ für den Block n in einem Frequenzband i in der Form

$$\hat{B}_{n,i} = \gamma_{n,i} \cdot \hat{B}_{n-1,i} + (1 - \gamma_{n,i}) \cdot \tilde{B}_{n,i}$$

$$\text{mit } \tilde{B}_{n,i} = \lambda_B \cdot \hat{B}_{n-1,i} + (1 - \lambda_B) \cdot S_{n,i}$$

erhalten wird,
wobei λ_B ein zwischen 0 und 1 liegender Vergessensfaktor ist, $\gamma_{n,i}$ der für den Block n im Frequenzband i bestimmte Grad der Stimmaktivität ist, und $S_{n,i}$ ein Mittelwert der Amplitude des Spektrums des Sprachsignals des Blocks n im Band i ist.

6. Verfahren nach Anspruch 5, bei dem das apriorisch rauschunterdrückte Signal $\hat{E}p_{n,i}$ bezüglich eines Blocks n und eines Frequenzbandes i die Form aufweist:

$$\hat{E}p_{n,i} = \max\left\{Hp_{n,i} \cdot S_{n,i}, \beta p_i \cdot \hat{B}_{n-\tau 1,i}\right\}$$

wobei $Hp_{n,i} = \dfrac{S_{n,i} - \alpha'_{n-\tau 1,i} \cdot \hat{B}_{n-\tau 1,i}}{S_{n-\tau 2,i}}$, $\tau 1$ eine ganze Zahl von mindestens gleich 1 ist, $\tau 2$ eine ganze Zahl von mindestens gleich 0 ist, $\alpha'_{n-\tau 1,i}$ ein für den Block $n-\tau 1$ und das Band i bestimmter Überbewertungskoeffizient ist, und βp_i ein positiver Koeffizient ist.

7. Verfahren nach einem der vorhergehenden Ansprüche, bei dem eine Langzeitschätzung ($\overline{E}_{n,i}$) der Energie des apriorisch rauschunterdrückten Signals ($\hat{E}p_{n,i}$)

berechnet wird und diese Langzeitschätzung mit einer an dem betreffenden Block berechneten, momentanen Schätzung (ba) dieser Energie verglichen wird, um den Grad der Stimmaktivität ($\gamma_{n,i}$) des Blocks zu erhalten.

8. Stimmaktivität-Erfassungseinrichtung mit einer zum Durchführen eines Verfahrens nach einem der vorhergehenden Ansprüche konzipierten Behandlungseinrichtung.

69803202.0
(98943998.9 / 1 016 071)
MATRAS NORTEL
COMMUNICATIONS

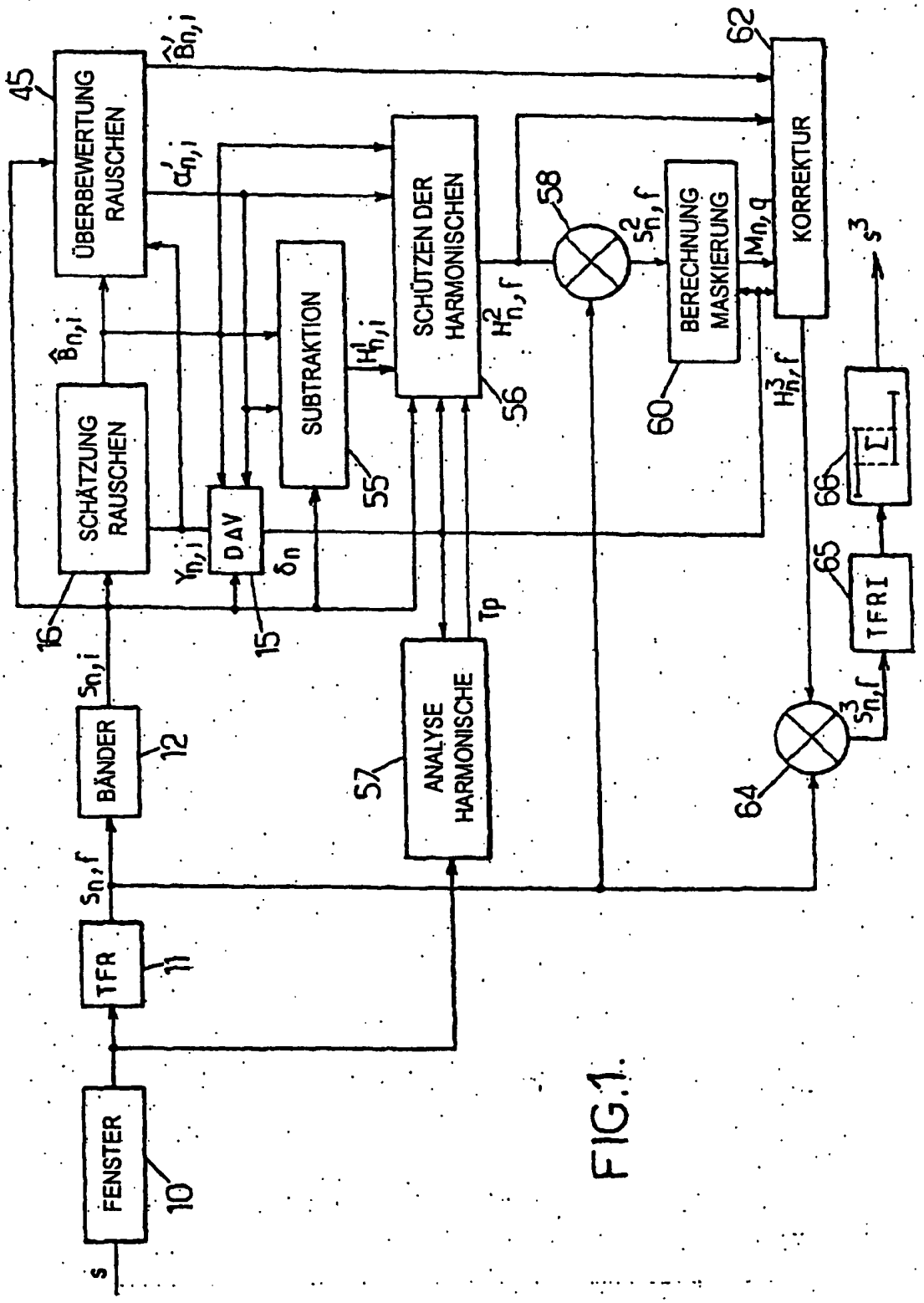


FIG.1.

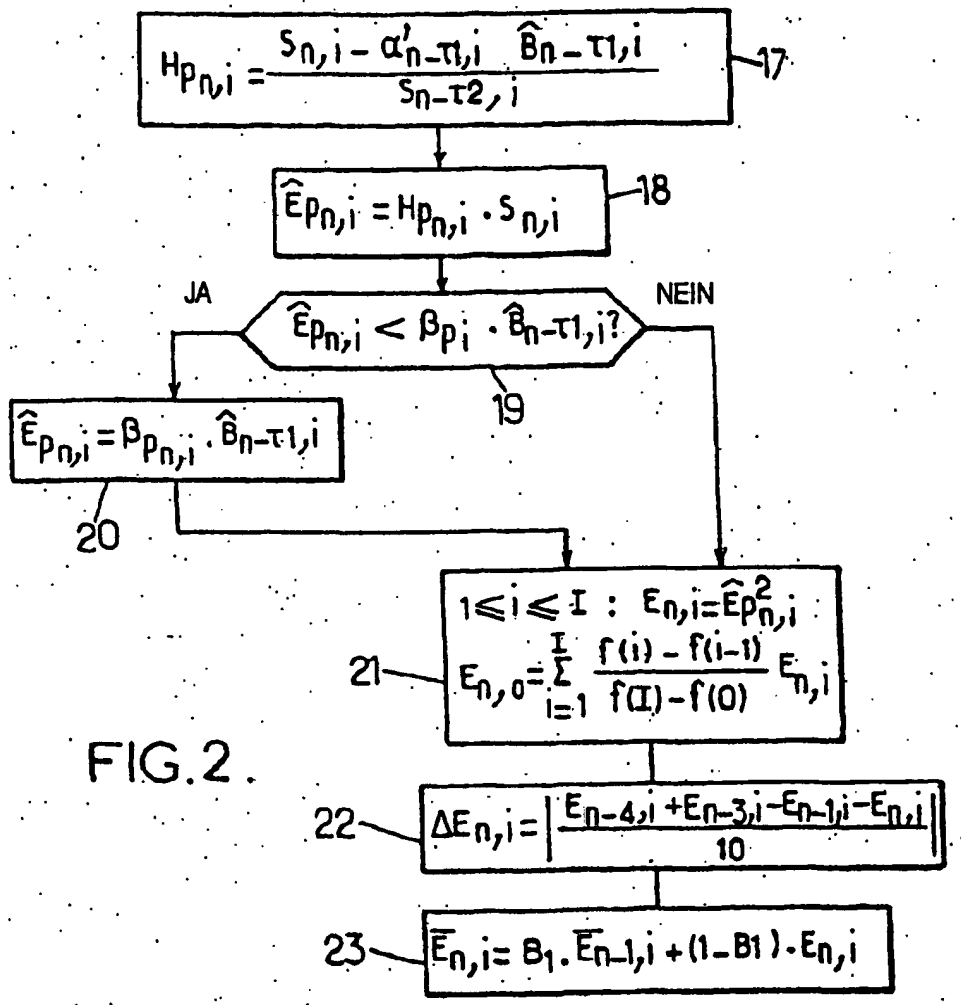

$Hp_{n,i} = \frac{S_{n,i} - \alpha'_{n-\tau1,i} \cdot \hat{B}_{n-\tau1,i}}{S_{n-\tau2,i}}$
17
$\hat{E}p_{n,i} = Hp_{n,i} \cdot S_{n,i}$
18
JA
NEIN
$\hat{E}p_{n,i} < \beta p_i \cdot \hat{B}_{n-\tau1,i}$?
19
$\hat{E}p_{n,i} = \beta p_{n,i} \cdot \hat{B}_{n-\tau1,i}$
20
$1 \leq i \leq I : E_{n,i} = \hat{E}p^2_{n,i}$
$E_{n,0} = \sum_{i=1}^{I} \frac{f(i) - f(i-1)}{f(I) - f(0)} E_{n,i}$
21
$\Delta E_{n,i} = \left| \frac{E_{n-4,i} + E_{n-3,i} - E_{n-1,i} - E_{n,i}}{10} \right|$
22
$\bar{E}_{n,i} = B_1 \cdot \bar{E}_{n-1,i} + (1 - B1) \cdot E_{n,i}$
23

FIG.2.

FIG.4.

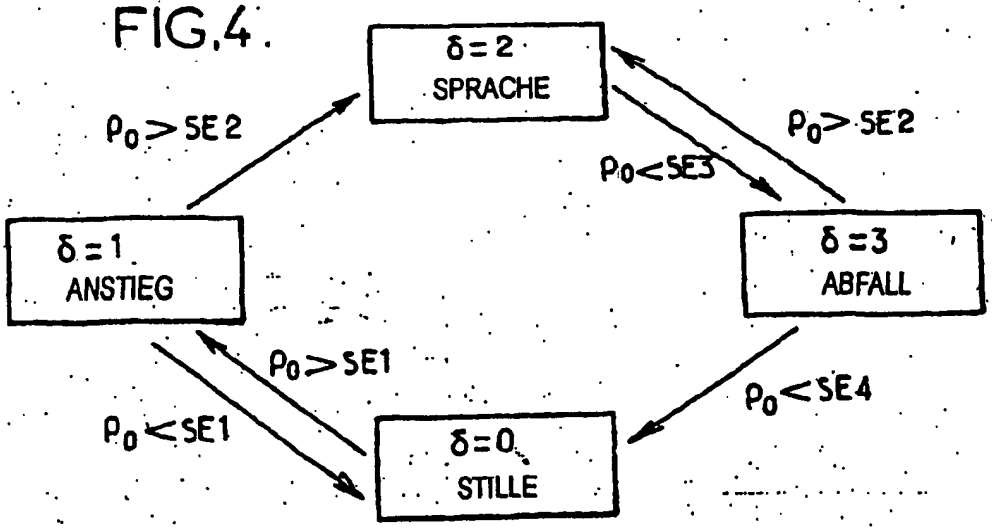
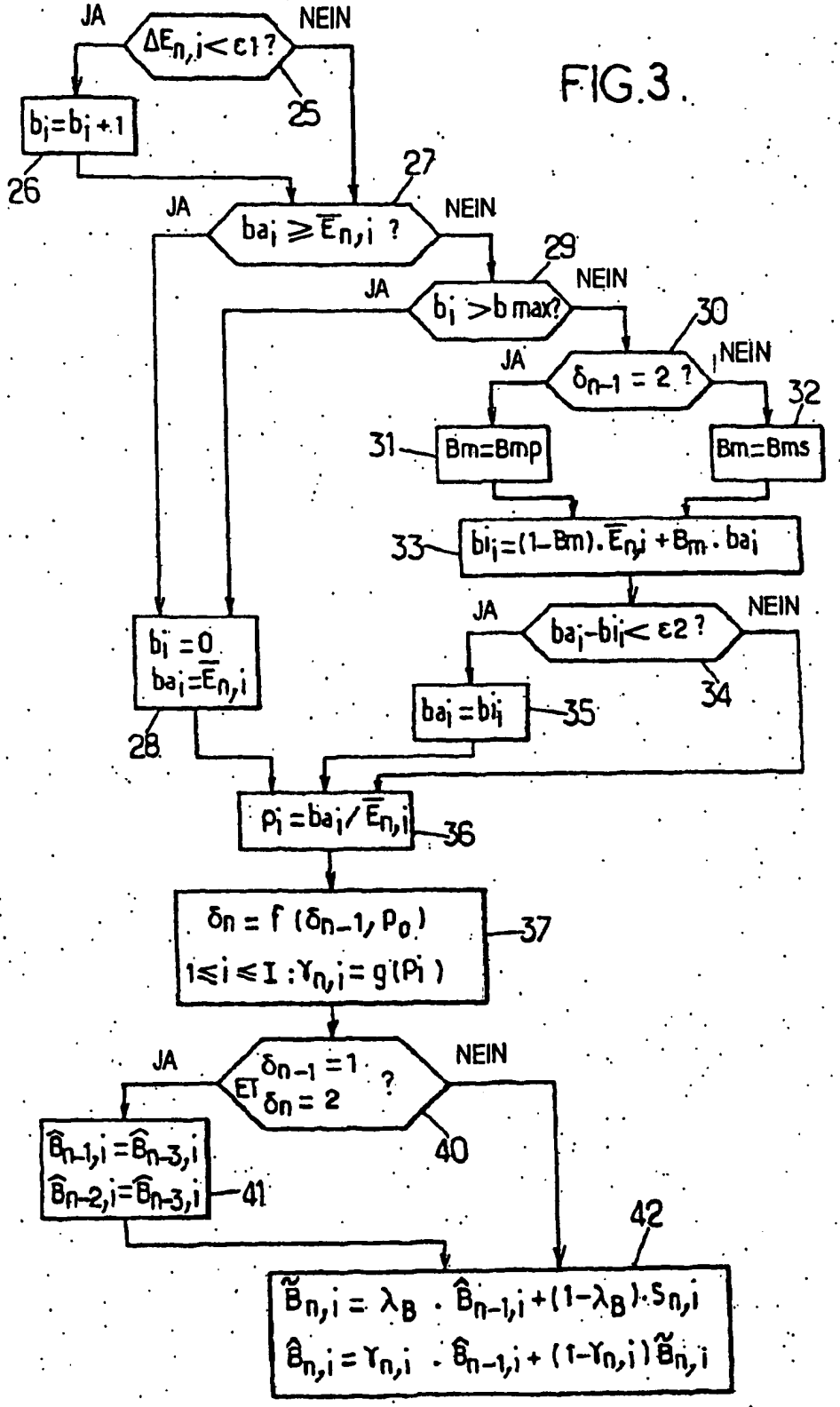

δ=2
SPRACHE
δ=1
ANSTIEG
δ=3
ABFALL
δ=0
STILLE
$p_0 > SE2$
$p_0 < SE3$
$p_0 > SE2$
$p_0 > SE1$
$p_0 < SE1$
$p_0 < SE4$

FIG.3.


JA
$\Delta E_{n,i} < \varepsilon 1$?
NEIN
25
$b_i = b_i + 1$
26
27
JA
$ba_i \geq \bar{E}_{n,i}$?
NEIN
29
JA
$b_i > b\,max$?
NEIN
30
JA
$\delta_{n-1} = 2$?
NEIN
32
31
$Bm = Bmp$
$Bm = Bms$
33
$bi_i = (1-Bm) \cdot \bar{E}_{n,i} + Bm \cdot ba_i$
JA
$ba_i - bi_i < \varepsilon 2$?
NEIN
34
$b_i = 0$
$ba_i = \bar{E}_{n,i}$
28
$ba_i = bi_i$
35
$p_i = ba_i / \bar{E}_{n,i}$
36
$\delta_n = f(\delta_{n-1}, p_0)$
$1 \leq i \leq I : \gamma_{n,i} = g(p_i)$
37
JA
$\delta_{n-1} = 1$ ET $\delta_n = 2$?
NEIN
40
$\hat{B}_{n-1,i} = \hat{B}_{n-3,i}$
$\hat{B}_{n-2,i} = \hat{B}_{n-3,i}$
41
42
$\tilde{B}_{n,i} = \lambda_B \cdot \hat{B}_{n-1,i} + (1-\lambda_B) \cdot S_{n,i}$
$\hat{B}_{n,i} = \gamma_{n,i} \cdot \hat{B}_{n-1,i} + (1-\gamma_{n,i}) \tilde{B}_{n,i}$

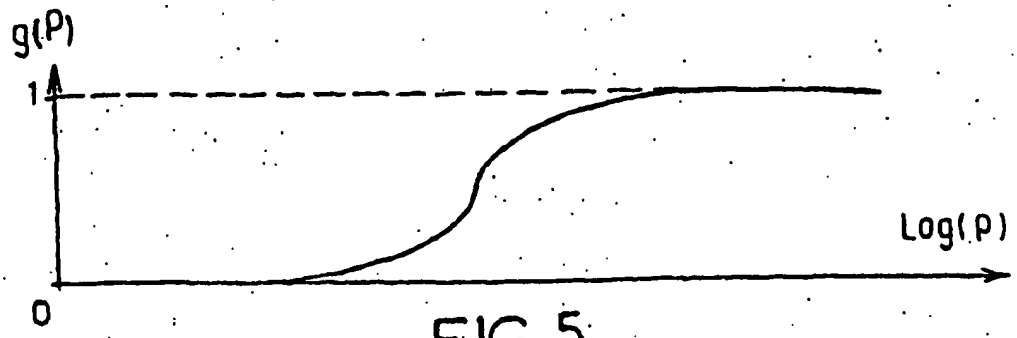

g(p)
1
0
Log(p)

FIG.5.

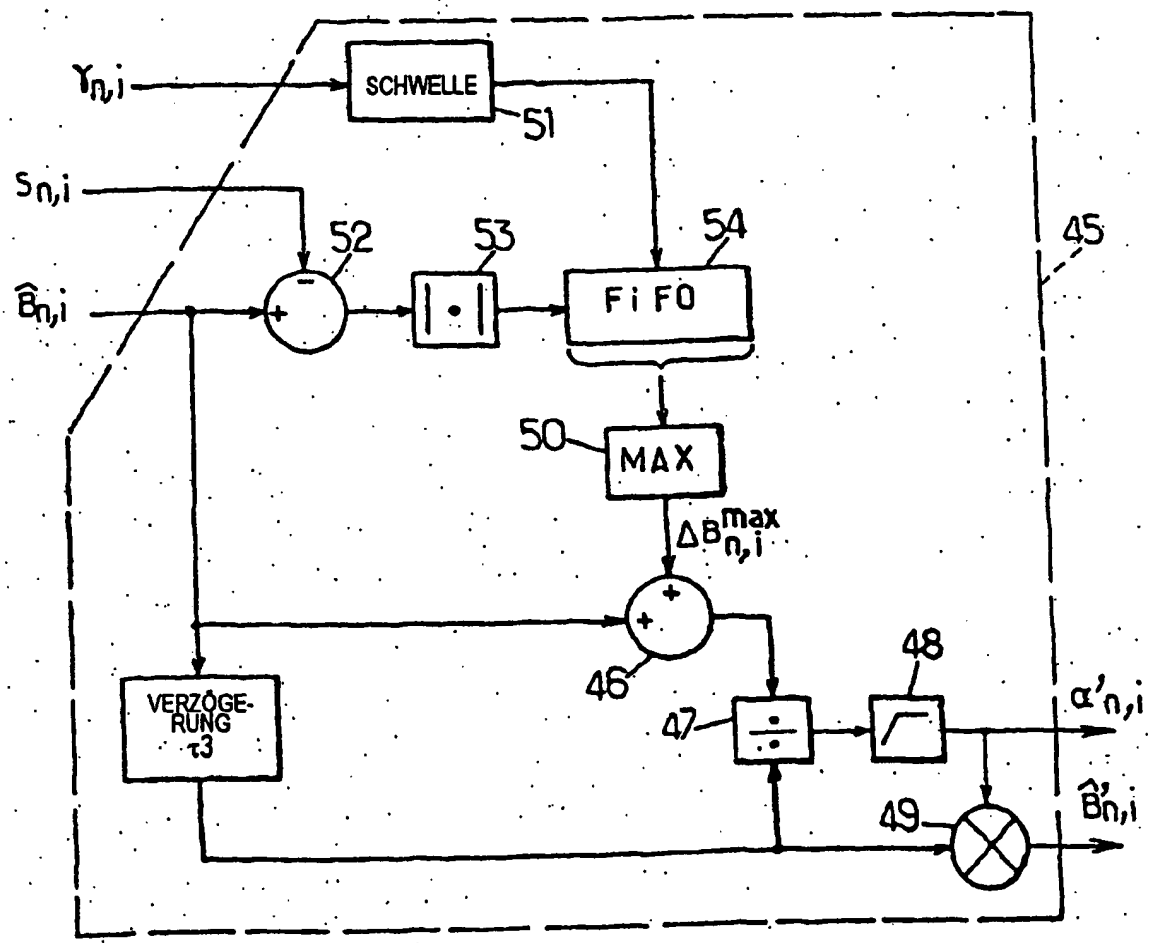

$\gamma_{n,i}$
SCHWELLE
51
$s_{n,i}$
$\hat{B}_{n,i}$
52
53
54
45
FIFO
50
MAX
$\Delta B^{max}_{n,i}$
46
47
48
VERZÖGE-
RUNG
τ3
$\alpha'_{n,i}$
$\hat{B}'_{n,i}$
49

FIG.6.

FIG.7.

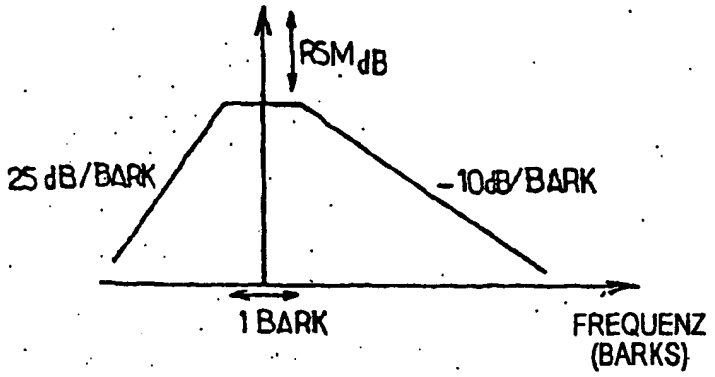
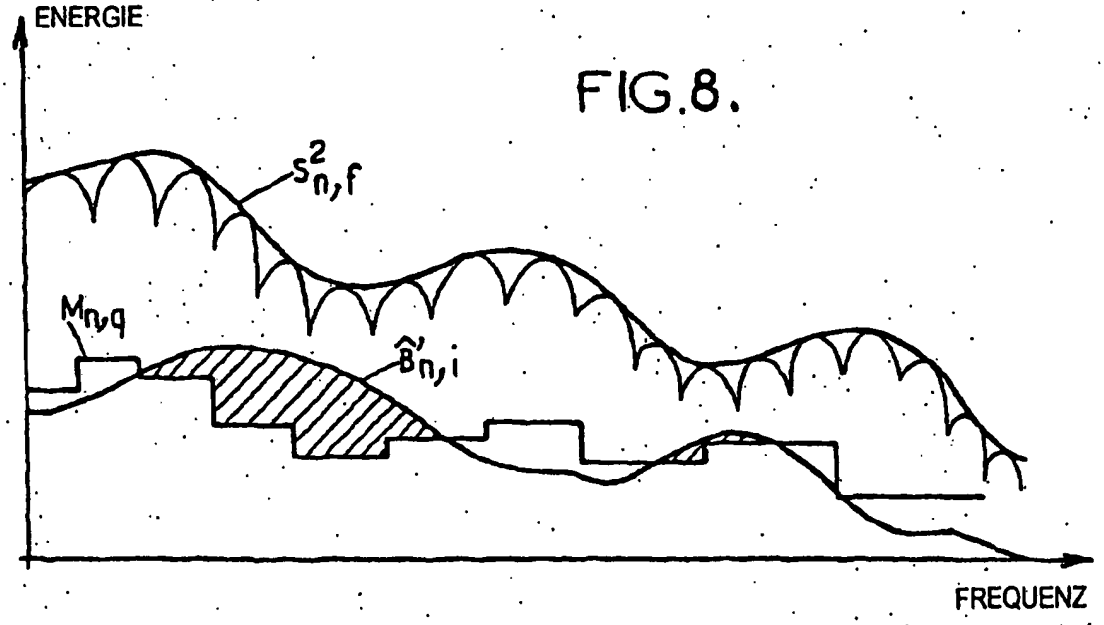

RSM dB
25 dB/BARK
-10dB/BARK
1 BARK
FREQUENZ
(BARKS)

FIG.8.


ENERGIE
$s^2_{n,f}$
$M_{n,q}$
$\hat{B}'_{n,i}$
FREQUENZ